DOCUMENT RESUME

ED 202 866 TM 810 282

AUTHOR Holloway, Robert E.
TITLE A Methodology for Assessing the Implementation of Educational Innovations: Analysis and Revision. Final Report.
INSTITUTION Columbia Univ., New York, N.Y. Teachers College.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C. Basic Skills Group.
PUB DATE Feb 81
GRANT NIE-G-78-0063
NOTE 47p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; *Instructional Design; Instructional Improvement; Instructional Innovation; *Multiple Regression Analysis; *Questionnaires; Secondary Education; *Secondary School Curriculum; Student Attitudes; Teacher Attitudes; *Test Reliability; *Test Validity
IDENTIFIERS *Personalized System of Instruction; Project Advance; Unit of Analysis Problems

ABSTRACT

An analysis of measures used in assessment of innovation fidelity was studied. The purpose was to contribute to the development of valid and reliable methods of measuring critical elements of an innovation and identifying levels of implementation. Questionnaires were developed for high school introductory courses in English, calculus, and psychology. Controlling for student aptitude and teacher characteristics in the calculus and psychology courses, no significant relationships were found when the student was used as the unit of analysis. Scales used on the questionnaires had sufficient reliability (internal consistency and stability) to suggest that they measure recognized factors in these courses. The English student questionnaire was developed using the five categories: materials, structure, behavior, knowledge, and values. This is in contrast to the characteristics of the calculus and psychology courses which were based on activities found in most Personalized System of Instruction (PSI) course designs. After administration, several factor solutions were generated; these were not judged as good fits to the five logical categories. Recommendations are: (1) analyses of implementation be done using classrooms, not students, as the unit of analysis; (2) categories of study be based on course-specific design factors; and (3) multiple outcome measures be used as dependent variables. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED202866

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

A METHODOLOGY FOR ASSESSING THE IMPLEMENTATION OF EDUCATIONAL INNOVATIONS: ANALYSIS AND REVISION

FINAL REPORT
SMALL GRANT (E)
78-0063

Submitted To

National Institue of Education
Basic Skills Group

By

Robert E. Holloway
Principal Investigator
Teachers College, Columbia
New York, N.Y.
10027

February 1981

TM 810 282

A METHODOLOGY FOR ASSESSING THE IMPLEMENTATION OF EDUCATIONAL INNOVATIONS: ANALYSIS AND REVISION

Robert E. Holloway
NIE Grant 78-0063

ABSTRACT

This study was an analysis of measures used in assessment of innovation fidelity. The purpose was to contribute to the development of valid and reliable methods of measuring critical elements of an innovation and identifying levels of implementation. Questionnaires were developed for introductory courses in English, calculus, and psychology. The student questionnaires for psychology and calculus were completed by 708 students in 40 classes in 33 high schools in the spring of 1978. The student questionnaire for English was completed by 314 students in 26 classes in 12 high schools in the fall of 1978.

Controlling for student aptitude and teacher characteristics in the calculus and psychology courses, no significant relationships were found when the student was used as the unit of analysis. Using the class as the unit of analysis 13% of the variance in total points (achievement) was predicted in the psychology course. The same procedure in the calculus course predicted 48% of the variance in the number of credit hours earned ($p < .05$). The scales used on the questionnaires had sufficient reliability (internal consistency and stability) to suggest that they measure recognized factors in these courses.

The English student questionnaire was developed using the five categories conceptualized by Fullan & Pomfret (1977): 1. materials, 2. structure, 3. behavior, 4. knowledge, and 5. values. This is in contrast to the characteristics of the calculus and psychology courses which were based on activities found in most Personalized System of Instruction (PSI) course designs. Content experts were given the Fullan & Pomfret categories and definitions of each. Items for the questionnaire were developed from course characteristics they identified. After administration several factor solutions were generated. The factor solutions were not judged as good fits to the five logical categories Fullan & Pomfret derived from the literature. Neither did other solutions appear interpretable. Subsequent analyses using dependent variables of achievement were not attempted.

Recommendations are that 1) analyses of implementation be done using classrooms as the unit of analysis rather than students, 2) categories of study be based on course-specific design factors rather than ostensibly generalizable categories, and 3) multiple outcome measures be used as dependent variables since variance within a single measure, such as grades, may not demonstrate relationships. The difficulty encountered in deriving low inference descriptive items for the questionnaires suggests further inquiry would be worthwhile only with innovations that are specified in considerable detail. Machine-based instruction is the most likely candidate for the immediate future.

TABLE OF CONTENTS

Purpose of the Study..... 1

Objectives..... 1

Data Analysis..... 1

Sample..... 6

Data Collection..... 9

Instrumentation.....11

Results.....13

TABLES

Number of Psychology Students..... 7

Number of Calculus Students..... 8

Data Collection Schedule.....10

Factor Loadings for PSI Questionnaire.....14

Factor Scales for PSI Questionnaire.....15

Summary of Regression of PSI Questionnaire and Points.....19

Summary of Regression of PSI Questionnaire and Credit.....20

Achievement Variance, Student and Class.....22

Achievement Variance, Student and Teacher.....22

FIGURES

Psychology Questionnaire Factor Scales.....16

Calculus Questionnaire Factor Scales.....17

UNPAGINATED
APPENDIX

References 4 Pages
Psychology Student Questionnaire 5 Pages
English Student Questionnaire 2 Pages
Abstracts of Presentations
Evaluating the Implementation of an Instructional Design...
American Educational Research Association, April 1979
The Development and Use of a Measure of Implementation...
Fifth National Conference on PSI, May 1979
A Methodology for Assessing Implementation...
International Communications Association, May 1980

Purpose

The purpose of this study is twofold: to develop a measure of the implementation of the PSI instructional design, and actually to use the measure to reveal the relationships between degrees of implementation of various dimensions of PSI and degrees of achievement on various criterion dimensions of specific PSI courses.

Objectives

The major objectives of this study are:

1) To develop teacher and student questionnaires with sufficient reliability and validity to be used as measures of implementation of PSI in two college-level courses taught in secondary schools;

2) to utilize the resulting questionnaires to provide process data which describe the implementation of the major dimensions of PSI in those courses;

3) to relate the measured differences in implementation to measured differences in student outcomes (achievement and attitude) while partialing out certain input variables (e.g., student aptitude and teacher enthusiasm, businesslike behavior, and clarity).

Data Analysis

The first task in the analysis of the data is to compute statistics which are descriptive of the implementation of the PSI

design in the two courses. The subprogram CONDESCRIPTIVE of the Statistical Package for the Social Studies (SPSS) (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975) will be used to determine the basic distributional characteristics of the data to be used in further analysis.

The second task is to reveal the effects of different degrees of implementation on the outcomes of the course. A multiple regression procedure (commonality analysis) recommended by Cooley and Lohnes (1976) and described by Kerlinger and Pedhazer (1973) will be used to answer this question. These procedures permit the researcher to estimate the percentage of criterion variance (e.g., student achievement) accounted for by predictor variables (e.g., student aptitude and degree of implementation) uniquely and in combinations. SPSS will also be used to process the data for this part of the analysis.

Cooley and Lohnes (1976) recommended that classes rather than students be the units of analysis in studies of this kind. However, Cooley (personal communication, April 4, 1978) and Lohnes (personal communication, April 28, 1978) recommended using students as the units of analysis in this particular study, primarily because a major portion of the process data, as well as all of the input and outcome data, will be collected at the student level.

Poynor's (1974) criticisms of using students as the units of analysis in evaluation research apply to comparative evaluations of a quasi-experimental nature. In such studies, the student can be the unit of analysis only if (1) the student is the

recipient of the treatment being investigated, and (2) the experimental conditions have been set up to balance teacher and classroom effects so that these influences are held constant. However, the present study is not in any sense a quasi-experimental evaluation of two competing educational treatments.

According to Poynor (1974), "[a] proper unit of analysis is the smallest source of data that is both logically and statistically defensible." The logical defense of students as the units of analysis for this study involves the following two points:

1. The primary unit of data collection is the student, and the data should be analyzed at the same level as it is collected whenever possible (Cooley & Leinhardt, 1975).
2. The primary question addressed by the data analysis is: What is the relationship of the implementation of PSI to academic achievement and attitude when the effects of certain input variables are eliminated? If it is found that variations in implementation do effect variations in achievement and attitude, a secondary analysis at the classroom level can be conducted to determine whether or not the variations in implementation are systematically related to the variable classes.

The major statistical defense of the units of analysis being students is that the degrees of freedom are severely limited when classes are used instead of students. Cooley and Leinhardt (1975) pointed out that at least 60 classrooms would be needed to properly analyze the data if six variable constructs are defined. This study will involve at least six variable constructs, and the

maximum number of classes available for data collection is 47.

In consideration of the above points, the unit of analysis for this research will be the student. Separate analyses will be conducted for the achievement and attitudes outcomes. The predictor domains for the achievement analysis will be input (as measured by high school grade point averages and the instructional rating forms) and process (as measured by the implementation questionnaires). The predictor variables for the attitude analysis will be input (as measured by the precourse attitude measures and the instructional rating forms) and process (as measured by the implementation questionnaires).

Importance of the Study

This study has immediate importance to the administrators, developers, faculty, and students involved in these courses, especially as it provides answers to the following two questions:

1. In what ways have teachers differed in their implementation of the PSI course design?

2. What are the effects of different degrees of implementation on the outcomes of the course?

Some possible decisions which could be influenced by the answers to these questions are as follows:

1. Faculty could be trained or otherwise influenced to bring their particular classroom practices into closer agreement with the intended practices.

2. Adaptational improvements found in particular classrooms could be disseminated to other teachers.

3. The course design and descriptions could be changed

to bring them closer in line with the realities of the classroom situation.

This study assumes that the primary goal of activities, such as instructional development and diffusion of educational innovations, is to increase or improve the outcomes of teaching and learning. The primary purpose of evaluation research such as the present study is to determine the extent to which the goal of these activities has been accomplished and the worth of the accomplishments.

In a more general way, this study is important to the improved practice of program evaluation. Charters and Jones (1973) criticize program evaluators for concentrating on developing and using appropriate outcome measures and elaborate research designs without describing and measuring the way in which their experimental and control situations differ. The "no significant differences" problem which has plagued the research and evaluation on the effectiveness of innovative programs and products in education (Allen, 1971) is at least partially the result of failing to describe and measure the unique attributes or qualities of the programs and products under study. As Shulman (1970) points out, we have measured aptitudes and outcomes by "micrometer" and treatments by "divining rods" (p. 374). This study attempts to discover the critical dimensions of a particular innovation (PSI) and to develop measures of the degree to which these dimensions are implemented in natural settings.

Sample

Approximately 900 students and 30 teachers will be surveyed and/or interviewed in this study. All subjects are either teaching or enrolled in one of two college freshmen-level courses which use the PSI course design. These courses, Psychology 205 and Self-Paced Calculus (see Appendix A for course descriptions), have undergone at least three years of design, evaluation, and redesign at Syracuse University. All subjects are participants in Syracuse University's Project Advance. (Tables 1 and 2)

Project Advance is a cooperative program between Syracuse University and New York State School districts. Selected courses, developed and implemented in the University by co-operating academic departments and the Center for Instructional Development, are piloted on campus and then offered for both high school and university credit in participating high schools as part of their regular school programs. Students are charged a modest overhead fee for the course and receive regular Syracuse University credit which is widely transferable to other colleges and universities throughout the country.

The courses are part of the regular teaching load of the high school teachers, who, as adjunct instructors, attend special university training workshops and seminars and teach the course under the supervision of university faculty. The grading standards for the course are identical both on and off campus.

Table 1

Number of Psychology Students Per School and Number of Psychology Students Per School in Final Sample

School Code	Number of Students Available	Number of Students in Final Sample	Percentage of Participation
A	10	8	80%
B	22	22	100%
C	22	19	86%
D[a]	46	40	87%
E	17	12	71%
F	38	30	79%
G	32	29	91%
H	11	10	91%
I	13	12	92%
J	28	27	96%
K	20	17	85%
L	16	13	81%
M	6	6	100%
N	20	20	100%
O	31	28	90%
P	11	8	73%
Q[b]	103	87	84%
R	34	14	41%
S	16	15	94%
T	34	30	88%
TOTAL	530	447	84%

[a]Had 2 classes with separate teachers

[b]Had 4 classes with same teacher

Table 2

Number of Calculus Students Per School and Number of Calculus Students Per School in Final Sample

School Code	Number of Students Available	Number of Students in Final Sample	Percentage of Participation
A	22	20	91%
B	17	15	88%
C	20	20	100%
D	9	0	0%
E	17	16	94%
F	14	0	0%
G	24	22	92%
H	22	16	73%
I	9	8	89%
J	32	11	34%
K	12	12	100%
L	16	15	94%
M	10	9	90%
N	16	16	100%
O	19	13	68%
P	27	23	85%
Q	9	6	67%
R	15	0	0%
S	40	39	98%
TOTAL	350	261	75%

Developed to meet a variety of needs expressed by high school superintendents, the project was first implemented during the 1973-74 academic year in six school districts. Over 400 students were enrolled in four of the five courses that were available. By the spring of 1978 the project had expanded to 80 schools in four states, New York, Massachusetts, Michigan, and New Jersey, and had an enrollment of approximately 4,500 students.

Though both the Psychology and Calculus courses employ the PSI design, they do differ in several respects. Psychology is a 3-credit-hour course and usually runs for one semester. Students successfully completing the course earn three hours of college credit. Calculus, as taught in the Project Advance high schools, continues throughout the academic year and students may earn from 2 to 12 credits. Calculus students must earn at least two credits per semester.

Data Collection

Stufflebeam et al., (1971) categorize evaluation data as either context, input, process, or product data. Eichelberger (1974) maintains that context data is more relevant to system-wide or state-wide evaluation than to evaluation of a particular program. Only input, process, and product data will be collected as part of this study.

Data collection for this study will be carried out as specified in Table 3.

TABLE 3

DATA COLLECTION SCHEDULE

Category	Description	Instrumentation	Time
Input	Student pre-course attitudes toward course and content	Adjective Rating Scale (Kelly & Greco, 1975)	February '78
Input	Student scholastic aptitudes as measured by high school grade-point-averages	School records	June '78
Input	Student ratings of teacher effectiveness, enthusiasm, businesslike behavior, and clarity	Student Instructional Report (Centra, 1972) (see Appendix E for copy.)	June '78
Process	Degrees of implementation of PSI course design	Student and teacher questionnaires	June '78
Product	Student Post-Course attitudes toward course and content	Adjective Rating Scale semantic differential	June '78
Product	Student Achievement as measured by (1) total points earned in course, (2) mid-term exams, and (3) standardized achievement tests	(1)(2) course records, and (3) CLEP	Throughout semester

Instrumentation

The major instruments to be developed for this study are the teacher and student questionnaires which will be used to measure degree of implementation. Fullan and Pomfret (1977) recommend that the use of questionnaires to measure implementation be explored because of their potential for reaching large samples of users. Fullan and Pomfret review several attempts to specify aspects of curricular innovations on questionnaires which seemed "fairly effective at measuring degree of implementation" (p. 366) based on observational and alternative instrument validity checks.

As described above, there are two perspectives from which implementation can be considered: the fidelity view (To what degree does the actual use of an innovation correspond to the intended user?) and the adaptation view (How does the innovation change in practice?). Fullan and Pomfret (1977) recommend that the fidelity perspective be adopted when studying the implementation of relatively explicit innovations, such as the PSI course design. An adaptation perspective is more appropriate when the innovation is less specific, e.g., team teaching, or when the innovation is in the early stages of development. For the purposes of this study, implementation will be measured from a fidelity perspective, though information will be sought which describes creative changes in the innovation.

The questionnaires will be developed according to the following process: (1) Information from a pilot study designed

to identify the critical dimensions of PSI courses will be combined with information from a review of the PSI literature to form the basis for questions. (2) Several questions will be written for each dimension of the course judged to be essential or at least desirable for the implementation of the PSI design. (3) The questionnaires will be reviewed by the course designers and a number of PSI authorities.

The kind of information sought by these questionnaires is primarily descriptive rather than judgmental. That is, the students and teachers will be asked to describe the actual activities which occurred in their courses rather than to make value judgments as to the worth of those activities. One type of reliability relevant to the questionnaires is the concept of stability of whether or not the respondents would give the same information on different occasions. The most common procedure for determining reliability of this sort is to actually administer the questionnaire or parts of it on two or more different occasions and to compare the consistency of the responses.

Results

Five research questions, based on the purposes of this study and the conclusions of the literature review, were addressed in a controlled correlational study. Before analyses related to the major questions were conducted, the student questionnaire was submitted to a principal components factor analysis with varimax rotation, and six factor scales were developed, labeled, and retained. Scores from the six scales (Lectures, Mastery/Difficulty, Self-Pacing, Written Materials, Feedback, and Peer Proctors) were also summed to constitute a Student Implementation Rating (SIR). The logical scales from the teacher questionnaire were reconstituted into six subscales supported through item analysis data. Scores from these six subscales (Mastery, Self-Pacing, Written Materials, Lectures as Motivators, Peer Proctors, and Feedback) were also summed to form a Teacher Implementation Rating (TIR). Standard scores (z-scores) from the Scholastic Aptitude Test (SAT) and grade-point-average were summed to form an indicator of student entering ability labeled Aptitude. Student ratings of teacher and course characteristics were submitted to an image factor analysis with varimax rotation. A five-item scale labeled TEACHCHAR was retained for use in further analyses as a measure of teacher characteristics.

Question 1 concerned the differences in implementa-

Table 4

Principal Components with Varimax Rotation Solution for Student PSI Implementation Questionnaire[a]

Item Number	Factor Number and Loadings 1	2	3	4	5	6	7	Communality
24	-.82[b]							.71
28	.78							.71
25	-.76							.60
26	.73							.55
2	.51						.30	.43
6		.81						.70
7		.75						.57
4		.63						.44
5		.47						.47
19			.83					.72
20			.75					.68
3			.63					.60
27			.38			.41		.42
29				.84				.75
30				.80				.73
17					.85			.77
18					.86			.78
8						.77		.65
9						.82		.70
1							.88	.80
Eigenvalue	3.21	2.45	1.81	1.70	1.41	1.23	1.04	
% of Variance	16.1	12.3	9.1	8.5	7.1	6.2	5.2	
Cumm. % of Variance	16.1	28.3	37.4	46.0	53.0	59.2	64.4	

[a] N = 598; 20 items

[b] Only loadings $\geq$ .30 are reported.

Table 4a

Factor Scales and Reliability Estimates for
Student PSI Implementation Questionnaire

Factor Label / Reliability[a]	Item Number	Question
Lectures	24	Lectures and demonstrations presented critical information which was covered in the unit tests.
α = .82	25	Lectures and demonstrations repeated the same information that was presented in the materials.
	26	Lectures and demonstrations presented new information not included in the unit tests.
	28	Lectures and demonstrations increased my interest in this course.
Mastery/ Difficulty	4	Scores necessary for passing module or unit tests were too high.
α = .64	6	There was too much content to be learned in each unit.
	7	The amount of content to be learned in the units was neither too much nor too little. It was just right.
Self-Pacing	3	Students in this course were permitted to take module or unit tests whenever they wanted to.
α = .64	19	I was permitted to work through the course at my own pace.
	20	The teacher set the pace at which I had to work through this course.
Written Materials	29	I could have passed this course by just studying the printed materials.
α = .80	30	The printed materials had to be accompanied by lectures and demonstrations in order for me to succeed.
Feedback	17	Usually my tests were scored and my results communicated to me.
α = .72	18	The length of time between taking tests and finding out the results was usually too long.
Peer Proctors	8	Students were used as proctors to grade tests, keep records, and/or tutor slower students.
α = .59	9	The teacher did the test scoring, grade keeping, and tutoring.

[a]Internal consistency estimate.

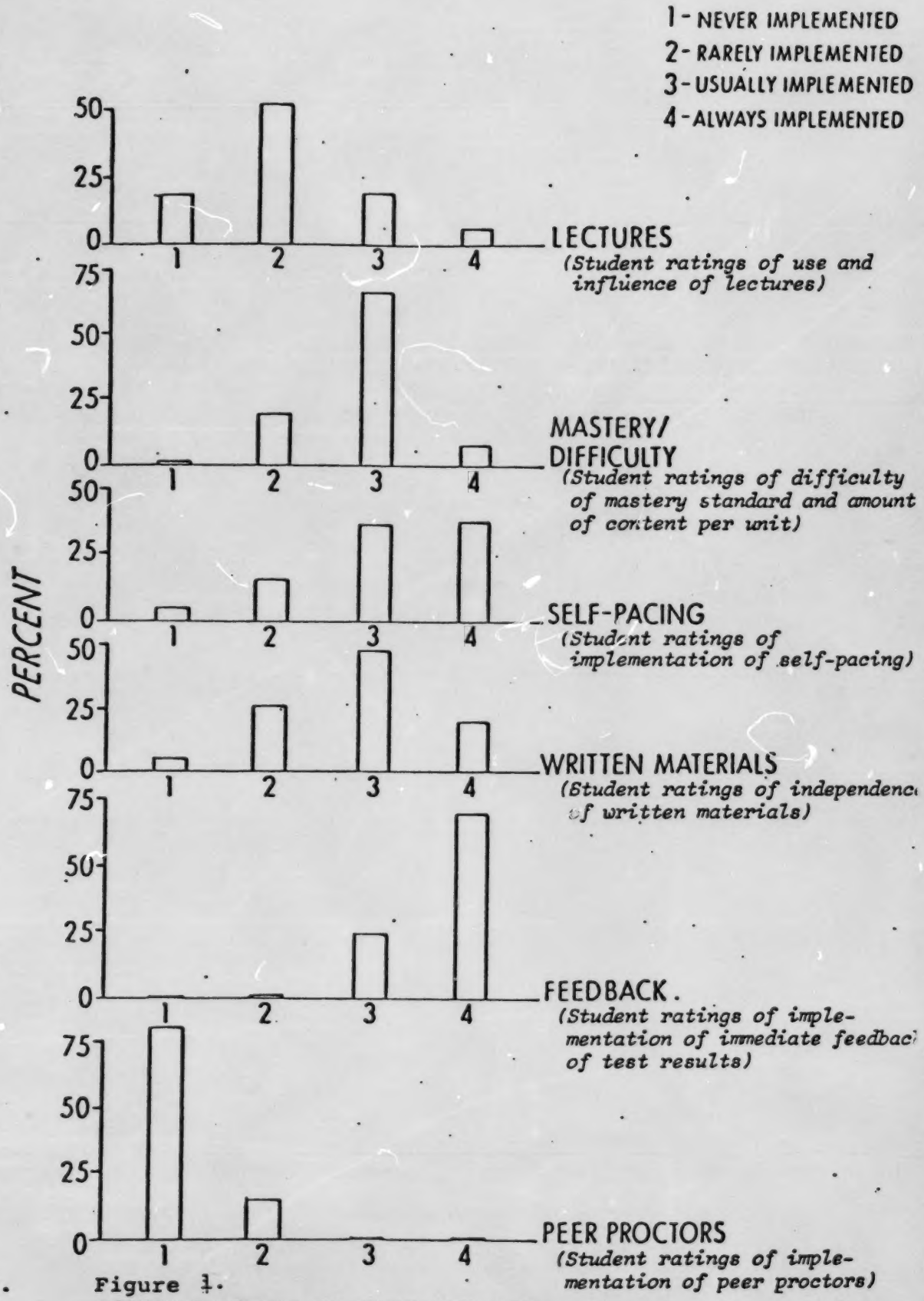
1 - NEVER IMPLEMENTED
2 - RARELY IMPLEMENTED
3 - USUALLY IMPLEMENTED
4 - ALWAYS IMPLEMENTED
PERCENT
LECTURES
(Student ratings of use and influence of lectures)
MASTERY/ DIFFICULTY
(Student ratings of difficulty of mastery standard and amount of content per unit)
SELF-PACING
(Student ratings of implementation of self-pacing)
WRITTEN MATERIALS
(Student ratings of independence of written materials)
FEEDBACK.
(Student ratings of implementation of immediate feedback of test results)
PEER PROCTORS
(Student ratings of implementation of peer proctors)

Figure 1.
PSYCHOLOGY FACTOR SCALES

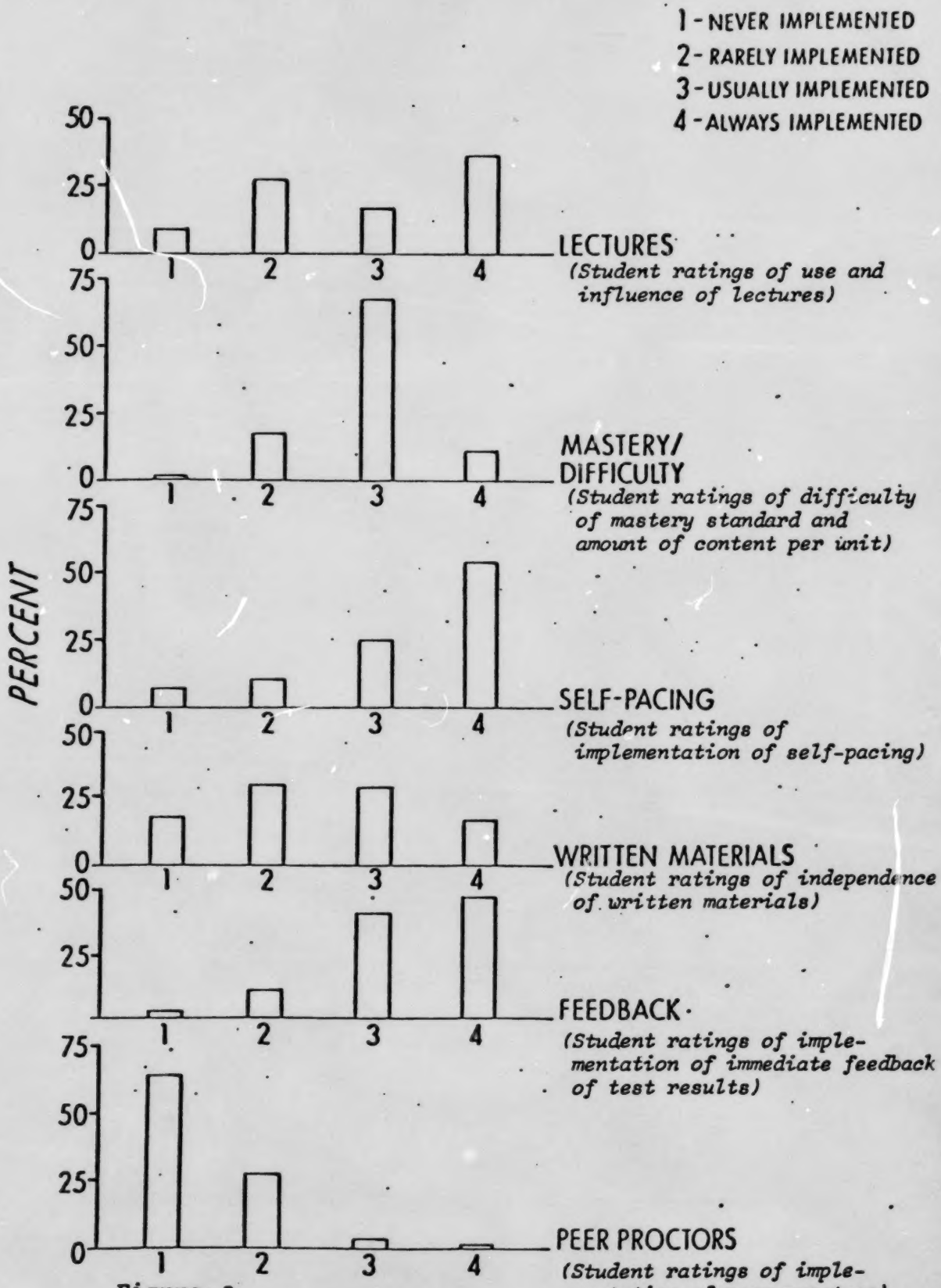

Figure 2.
CALCULUS FACTOR SCALES

tion of the major dimensions of PSI as measured by the student questionnaire. The differences among classes on the total and factor scale scores of the SIR and the low intercorrelations among the factor scales of the SIR supported the conclusion that the student questionnaire detected differences in the implementation of PSI in the psychology and calculus classes examined in this study.

Question 2 concerned the differences in implementation of the major dimensions of PSI as measured by the teacher questionnaire. The differences between the ratings of the subscales of the TIR given by the psychology and calculus teachers, the differences among classes on total and subscale scores of the TIR, and the low intercorrelations among the subscale scores supported the claim that the teacher questionnaire detected differences in the implementation of PSI.

Commonality analysis and partial correlation procedures were used for question 3 to determine the amount of variance in student achievement accounted for by variation in the implementation of PSI as measured by the SIR when the effects of Aptitude and TEACHCHAR were controlled. Using the student as the unit of analysis, SIR predicted 1% of the variance in Total Points in the psychology course and 7% of the variance in NUMHOURS in the calculus course ($p < .01$). Using the class as the unit of analysis, SIR predicted 13% of the variance in Total Points and 48% of the variance in NUMHOURS ($p < .05$). The Mastery/Difficulty

Table 5

Summary of Multiple Regression Analysis[a] for PSI Subscales of Student Questionnaire Against Total Points

(N = 349)

Factor Variable[b]	F value[c] for amount of variance accounted for by addition of each variable	Cummulative % of variance of Total Points accounted for by addition of each variable
1 - Aptitude	53.155*	18.1%
2 - TEACHCHAR	1.813	20.4%
3 - Mastery/ Difficulty	17.169*	24.2%
4 - Self-Pacing	1.522	24.6%
5 - Peer Proctors	.521	24.7%
6 - Feedback	.326	24.8%
7 - Lectures	.012	24.8%
8 - Written Materials	.009	24.8%

[a] Total $R^2_{y.12345678} = .4979$

[b] Aptitude and TEACHCHAR were forced to enter first in this analysis. Remaining variables entered in stepwise regression.

[c] Degrees of freedom = 8,340

* $p < .01$

Table 6

Summary of Multiple Regression Analysis[a] for PSI Subscales of Student Questionnaire Against NUMHOURS

(N = 181)

Predictor Variable[b]	F value[c] for amount of variance accounted for by addition of each variable	Cummulative % of variance of NUMHOURS accounted for by addition of each variable
1 - Aptitude	11.547*	7.7%
2 - TEACHCHAR	9.497*	13.3%
3 - Lectures	4.508*	18.8%
4 - Written Materials	1.480	19.5%
5 - Feedback	1.547	20.2%
6 - Peer Proctors	.327	20.3%
7 - Mastery/ Difficulty	.257	20.4%
8 - Self-Pacing	.139	20.5%

[a] Total $R^2y.12345678 = .4530$

[b] Aptitude and TEACHCHAR were forced to enter first in this analysis. Remaining variables entered in stepwise regression.

[c] Degrees of freedom = 8,172

*$p < .01$

factor scale predicted 4% of the variance in Total Points and the Lectures scale predicted 6% of the variance in NUMHOURS with the student as the unit of analysis ($p < .01$). Using the class as the unit of analysis, the Self-Pacing and Peer Proctors factor scales predicted respectively 14% and 16% of the variance in Total Points, and the Lectures, Written Materials, and Feedback scales predicted respectively 55%, 26%, and 27% of the variance in NUMHOURS ($p < .05$). The partial correlations used to analyze the data when the class was the unit of analysis do not take into account the intercorrelations of the factor scales. Thus these results are suspect.

Partial correlation analysis was used in question 4 to determine the amount of variance in student achievement accounted for by variation in the implementation of PSI as measured by the TIR when the effects of Aptitude and TEACHCHAR were controlled. No statistically significant relationships ($p < .05$) were detected between TIR and Total Points in the Psychology classes. In calculus, TIR predicted 52% of the variance in NUMHOURS when the effects of Aptitude and TEACHCHAR were partialed out ($p < .05$). The subscales Mastery, Written Materials, and Peer Proctors predicted statistically significant amounts of variance in NUMHOURS (29%, 32%, and 36% respectively) ($p < .05$). However, these last results are possibly inflated since they capitalize on the interscale correlations on the TIR.

The predictive validity of the student questionnaire was investigated by estimating the predictive power of the SIR against student outcomes (achievement and attitude) when the effects of student ability and teacher characteristics were controlled. The percentages of criterion variance in achievement predicted by the SIR in the psychology and calculus courses with both student and class as the unit of analysis are summarized in Table 7. The SIR did not predict a meaningful amount of variance in attitude. The SIR predicted significant amounts of variance in student achievement when

Table 7

Percentages of Achievement Variance Accounted for by SIR When the Effects of Aptitude and TEACHCHAR Are Controlled

Unit of Analysis	Psychology (Total Points)	Calculus (NUMHOURS)
Student	1%**	7%**
Class	13%**	48%*

*p < .05
**p < .01

Table 8

Percentages of Achievement Variance Accounted for by Implementation Questionnaires When the Effects of Aptitude and TEACHCHAR are Controlled
Unit of Analysis = Class

Implementation Measure	Psychology (Total Points)	Calculus (NUMHOURS)
SIR	13%*	48%*
TIR	4%	52%*

*p < .05

Aptitude and TEACHCHAR were controlled ($p < .05$). The predictive validity of the TIR was supported in the calculus sample, but not in the psychology sample. The convergent validity of the SIR and TIR was moderately supported by the finding that the Spearman rank-difference correlation between the two measures was .59 ($p < .05$). The divergent validities of the SIR and TIR were apparently contradicted by the significant ($p < .05$) correlations between the SIR and Aptitude and TIR and TEACHCHAR in the psychology sample.

Question 5 concerned the relative utility of the student and teacher questionnaires. The SIR was judged more effective than the TIR in predicting statistically significant amounts of achievement variance in the psychology and calculus courses. Neither questionnaire was effective in determining the relative contributions of the major dimensions of PSI to the prediction of student achievement. Given the greater effectiveness of the student questionnaire, the relatively greater costs of administering it appear justifiable.

END-PART I

References

Allen, W. H. Instructional media research: Past, present, and future. AV Communication Review, 1971, 19, 5-18.

Anastasi, Anne. Psychological testing (3rd ed.). New York: MacMillan, 1968.

Armor, David J. Theta reliability and factor scaling. In Herbert L. Costner (Ed.). Sociological Methodology 1973-1974. San Francisco, Calif.: Jossey-Bass, 1974.

Bond, G. L., & Dykstra, R. The cooperative research program in first-grade reading instruction. Reading Research Quarterly, 1967, 2 (4), 5-142.

Bruner, J. S. Toward a theory of instruction. Cambridge, Mass.: Belnap Press, 1966.

Carroll, J. B. Learning and individual differences. Columbus, Ohio: Charles E. Merrill, 1967.

Centra, J. A. The student instructional report: Its development and uses. SIR Report #1. Princeton, N.J.: Educational Testing Service, 1972.

Charters, W. W., & Jones, J. On the risk of appraising non-events in program evaluation. Educational Researcher, 1973, 2 (11).

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. Equality of educational opportunity. Washington, D.C.: U.S. Government Printing Office, 1966.

Cooley, W. W. & Lohnes, P. R. Evaluation research in education: Theory, principles, and practice. New York: Irvington, 1976.

Eichelberger, R. T. Evaluating ongoing instructional programs. In Gary D. Borich (Ed.). Evaluating educational programs and products. Englewood Cliffs, N.J.: Educational Technology Publications, 1974.

Fraley, L. E., & Vargas, E. A. (Eds.). Proceedings of the third national conference on behavioral research and technology in higher education. Atlanta, Ga.: Society for Behavioral Technology and Engineering, 1976.

Fullan, M., & Pomfret, A. Research on curriculum and instruction implementation. Review of Educational Research, 1977, 47, 335-397.

Gagne, R. M. The conditions of learning (2nd ed.). New York: Holt, Rinehart and Winston, 1970.

Gephart, William J. Problems in measuring the degree of implementation of an innovation. Paper presented at the annual convention of the American Educational Research Association, San Francisco, April 1976.

Glaser, R. Adapting the elementary school curriculum to individual performance. In Proceedings of the 1967 Conference on Testing Problems. Princeton, N.J.: Educational Testing Service, 1968.

Glass, G.V. Two generations of evaluation models. In P. A. Taylor & D. M. Cowley (Eds.), Readings in curriculum evaluation. Dubuque, Iowa: Wm. C. Brown, 1972.

Gross, N., Giacquinta, J. B., & Bernstein, M. Implementing organizational innovation: A sociological analysis of planned educational change. New York: Basic Books, 1971.

Ingle, Robert B. Evaluation parameters: As a class of activities; As applied to programs. The CEDR Quarterly, 1976, 9(3), 7-10.

Johnston, J. M. Behavior research and technology in higher education. Springfield, Illinois: Charles C. Thomas, 1975.

Johnston, J. M., & O'Neill, G. W. Proceedings of the second national conference on research and technology in college and university teaching. Gainesville, Fla.: James M. Johnston, 1975.

Journal of Personalized Instruction. Washington, D.C.: Center for Personalized Instruction, 1975-1978.

Keller, F. S. "Goodby, teacher..." Journal of Applied Behavioral Analysis, 1968, 1, 79-89.

Keller, F. S., & Sherman, J. G. The Keller Plan handbook. Menlo Park, Calif.: W. A. Benjamin, 1974.

Kelly, Edward, & Greco, Thomas. Manual for the Adjective Rating Scale. Syracuse, N.Y.: Center for Instructional Development, Syracuse University, 1975.

Kerlinger, F. N., & Pedhazur, E. J. Multiple regression in behavioral research. New York: Holt, Rinehart and Winston, 1973.

Klein, S.P. Evaluation workshop I: An orientation. Monterey Park. Calif.: Center for the Study of Evaluation, UCLA, CTB McGraw-Hill, 1971.

Kulik, J. A., Kulik, C. L., & Carmichael, K. The Keller plan in science teaching. Science, 1974, 183, 379-383.

Leinhardt, G. Program evaluation: An empirical study of individualized instruction. American Educational Research Journal, 1977, 14, 277-293.

Lindvall, C.M., Cox, R.C. Evaluation as a tool in curriculum development: The IPI evaluation program. Chicago: Rand McNally, 1970.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. Statistical Package for the Social Studies (2nd. ed.). New York: McGraw-Hill, 1975.

Personalized System of Instruction Newsletter . Washington, D.C.: Center for Personalized Instruction, 1971-78.

Results of Follow Through evaluation disappointing but useful, OE says. Report on Educational Research, August 24, 1977, p. 9.

Robin, A. L. Behavioral instruction in the college classroom. Review of Educational Research, 1976, 46, 313-354.

Rosenshine, Barak & Furst, Norma. The use of direct observation to study teaching. In Robert M. W. Travers (Ed.). Second handbook of research on teaching. Chicago, Il.: Rand McNally, 1973.

Ruskin, T. S., & Bono, S. F. Proceedings of the first national conference. Washington, D.C.: Center for Personalized Instruction, 1974.

Ryan, B. A. PSI: Keller's personalized system of instruction: An appraisal. Washington, D. C.: American Psychological Association, 1974. (monograph)

Scheffe, H. The analysis of variance. New York: Wiley, 1959.

Sherman, J. G. (Ed.). Grants kick off new center. Personalized System of Instruction Newsletter, 10, October, 1973.

Sherman, J.G. (Ed.). Personalized system of instruction: 41 germinal papers. Menlo Park, Calif.: W. A. Benjamin, 1974.

Shulman, L.S. Reconstruction of educational research. Review of Educational Research, 1970, 40, 371-396.

Skinner, B. F. The science of learning and the art of teaching. Harvard Educational Review, 1954, 24, 86-97.

Skinner, B. F. Teaching machines. Science, 1958, 128, 969-977.

Stake, R. The countenance of educational evaluation. Teachers College Record, 1967, 68, 523-540.

Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond, R. L., Merriman, H. O., & Provus, M. M. Educational evaluation and decision making. Itasca, Ill.: F. E. Peacock, 1971.

Wispe, L. G. Evaluating section teaching methods in an introductory course. Journal of Educational Research, 1951, 45, 161-186.

Psychology Student Questionnaire

NAME _____

SCHOOL _____ SOCIAL SECURITY NUMBER _____

COURSE _____ DATE _____

PROJECT ADVANCE EVALUATION
SPRING 1978

Instructions: Please be sure to fill in your name, Social Security number, school, date, and course in the spaces above.

The purpose of this questionnaire is to find out what your class was like during the past semester. This is not a "test." Please give honest, frank opinions about the Project Advance course you are now attending. Your responses will be kept confidential.

Some of the questions ask you to describe particular activities in your class. Others ask for your opinion of those activities in your class.

In responding to each question go through the following steps:

1. Read the statement carefully.
2. Read the four responses below the statement and place an X or ✓ next to the word or phrase which best describes the activities in your class or expresses your opinion about the class activities. For example, if the statement is:

   Attendance was optional in this course.

   and attendance was required, you would check the *"NEVER"* response in the list below the statement.

   *ALWAYS*__ *USUALLY*__ *RARELY*__ *NEVER*✓
3. Answer every item. Please feel free to write comments anywhere on this questionnaire.

N.B. The words "module" and "unit" are used synonymously in this questionnaire.

© Copyright Center for Instructional Development, Syracuse University, 1978

ONWARD!!!

1. Students had to score at least 80% on module or unit tests in order to pass the tests.

   ALWAYS ___ USUALLY ___ RARELY___ NEVER___

2. Students had to pass the test on the preceding module or unit before going on to the next module.

   ALWAYS___ USUALLY___ RARELY___ NEVER___

3. Students in this course were permitted to take module or unit tests whenever they wanted to.

   ALWAYS___ USUALLY___ RARELY___ NEVER___

4. Scores necessary for passing module or unit tests were too high.

   STRONGLY AGREE___ AGREE___ DISAGREE___ STRONGLY DISAGREE___

5. The number of opportunities to take tests was adequate.

   STRONGLY AGREE___ AGREE___ DISAGREE___ STRONGLY DISAGREE___

6. There was too much content to be learned in each unit.

   STRONGLY AGREE___ AGREE___ DISAGREE___ STRONGLY DISAGREE___

7. The amount of content to be learned in the units was neither too much nor too little. It was just right.

   STRONGLY AGREE___ AGREE___ DISAGREE___ STRONGLY DISAGREE___

8. Students were used as proctors to grade tests, keep records, and/or tutor slower students.

   ALWAYS___ USUALLY___ RARELY___ NEVER___

9. The teacher did the test scoring, record keeping, and tutoring.

   ALWAYS___ USUALLY___ RARELY___ NEVER___

10. The teacher was available to assist individual students with their work during regular scheduled Project Advance classes.

    ALWAYS___ USUALLY___ RARELY___ NEVER___

11. The teacher was available to assist individual students with their work at times other than the regular scheduled Project Advance classes.

    ALWAYS___ USUALLY___ RARELY___ NEVER___

12. The assistance I received from the teacher helped me to complete my course work.

    ALWAYS___ USUALLY___ RARELY___ NEVER___

13. I read the objectives in the beginning of each module or unit.

ALWAYS ___ USUALLY ___ RARELY ___ NEVER ___

14. The objectives let me know what I had to learn in order to pass the unit tests.

ALWAYS ___ USUALLY___ RARELY ___ NEVER ___

15. I answered the study guide questions before taking module tests.

ALWAYS ___ USUALLY ___ RARELY ___ NEVER ___

16. The study guide questions helped me to pass the unit tests.

ALWAYS ___ USUALLY ___ RARELY ___ NEVER ___

17. Usually my tests were scored and my results communicated to me

AS SOON AS I HANDED IN THE TEST ___
BY THE NEXT SCHEDULED CLASS ___
WITHIN A WEEK ___
AFTER MORE THAN A WEEK ___

18. The length of time between taking tests and finding out the results was usually too long.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

19. I was permitted to work through the course at my own pace.

ALWAYS ___ USUALLY ___ RARELY ___ NEVER ___

20. The teacher set the pace at which I had to work through the course.

ALWAYS ___ USUALLY ___ RARELY ___ NEVER ___

(You do not need to answer questions 21 and 22 if you checked RARELY *or* NEVER *for question 20.)*

21. The pace was just right.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

22. The pace was too slow.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

23. The initial tests for a module and the retests for that module were of equal difficulty. *(Skip this question if you took no retests.)*

ALWAYS ___ USUALLY ___ RARELY ___ NEVER ___

24. Lectures and demonstrations presented critical information which was covered in the unit tests.

ALWAYS ___ USUALLY ___ RARELY ___ NEVER ___

25. Lectures and demonstrations repeated the same information that was presented in the printed materials.

ALWAYS ___ USUALLY ___ RARELY ___ NEVER ___

26. Lectures and demonstrations presented new information that was not included in the unit tests.

ALWAYS ___ USUALLY ___ RARELY ___ NEVER ___

27. Attendance at lectures and demonstrations was optional.

ALWAYS ___ USUALLY ___ RARELY ___ NEVER ___

28. Lectures and demonstrations increased my interest in this course.

ALWAYS ___ USUALLY ___ RARELY ___ NEVER ___

29. I could have passed this course by just studying the printed materials.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

30. The printed materials had to be accompanied by lectures and demonstrations in order for me to succeed.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

31. The number of points required to get an "A" in this course was too high.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

32. I could choose the grade I wanted to work toward: "A", "B", "C", or "D".

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

33. This teacher enjoys teaching.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

34. In this course the tests correspond well to the study materials.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

35. This teacher explains complicated ideas and relationships clearly.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

36. I have used what I am learning in this course in other courses I am taking.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE___

37. We spent a lot of time fooling around in this class.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

38. In this class we learn things that are very practical.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

39. This teacher is clear about what I am expected to learn.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

40. In this class what I am tested on corresponds well to what I have a chance to learn.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

41. This teacher gets right down to business in this class.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

42. This teacher is enthusiastic about teaching.

STRONGLY AGREE ___ AGREE ___ DISAGREE ___ STRONGLY DISAGREE ___

For items 43 through 65 please use the following code:

1 = EXTREMELY 2 = VERY 3 = SLIGHTLY 4 = NOT AT ALL

I FOUND THIS PROJECT ADVANCE COURSE TO BE:

43. ___ INTERESTING	48. ___ GOOD	54. ___ DULL	60. ___ ENJOYABLE
44. ___ BORING	49. ___ STIMULATING	55. ___ CHALLENGING	61. ___ ENLIGHTENING
45. ___ RELEVANT	50. ___ IRRELEVANT	56. ___ A WASTE	62. ___ EXCITING
46. ___ INFORMATIVE	51. ___ WORTHWHILE	57. ___ PRACTICAL	63. ___ REWARDING
47. ___ DIFFICULT	52. ___ VALUABLE	58. ___ DEMANDING	64. ___ PROVOCATIVE
	53. ___ NECESSARY	59. ___ DIFFERENT	65. ___ USELESS

66. Give this course a grade. A ___ B ___ C ___ D ___

67. Give the teacher a grade. A ___ B ___ C ___ D ___

68. Give the content of this course a grade. A ___ B ___ C ___ D ___

(The last three questions ask for specific information about you. As noted on the first page, all your answers will be kept confidential.)

69. What is your sex? Male ___ Female ___

70. What is your overall high school grade average?

A ___ B ___ C ___ D ___

71. What are your Scholastic Aptitude Test (SAT) scores?*

Verbal Score _____ Math Score _____

**If you have not taken the SAT exams, but have taken the PSAT exams, please report the PSAT scores.*

NAME _____

SOCIAL SECURITY NUMBER (_____)

Please return by May 25.

ENGLISH POETRY COURSE

We would appreciate your help in finding out how students who have taken the English Poetry course feel about it. Your response will not be recorded for your teacher nor will the results affect the course as it will be taught next year. We do plan, however to determine which parts of the course are important to you and which are not and, through working with all teachers, suggest how the prescribed elements are valued by students.

Listed below are statements which might describe your poetry course. Simply circle the degree to which each statement accurately describes the course in your judgment. Specific comments about things that you found beneficial or not, whether or not they are mentioned in the statements below, will help us discover additional ways in which the course may be changed for future students.

	Disagree Strongly	Disagree	Neutral	Agree	Agree Strongly
EXAMPLE: The poetry course was the same as the other English mini-courses.	DS	D	N	A	AS

In the example, the respondent felt the statement was somewhat accurate but had reservations. Neutral ("N") can mean you have no strong reaction to the descriptive statement or that you are unsure or ambivalent.

1. An objective of the course was to teach me to identify the formal elements of poetry.	DS	D	N	A	AS
2. The identification of formal elements of poetry was a valuable objective.	DS	D	N	A	AS
3. An objective of the poetry course was to increase my appreciation of poetry.	DS	D	N	A	AS
4. Increasing my appreciation of poetry was a worthwhile objective.	DS	D	N	A	AS
5. An objective of the poetry course was to develop my writing skills.	DS	D	N	A	AS
6. The development of my writing skills was an important objective.	DS	D	N	A	AS
7. An objective of the poetry course was to develop literary insights.	DS	D	N	A	AS
8. The development of literary insights was a valuable objective.	DS	D	N	A	AS
9. What other objectives did you find the course met or should have incorporated?					

	Disagree Strongly	Disagree	Neutral	Agree	Agree Strongly
10. Being able to understand poetry is important.	DS	D	N	A	AS
11. The out-of-class papers I wrote were based on poems we studied in class.	DS	D	N	A	AS
12. I had one or more conferences on a one-to-one basis with my poetry teacher.	DS	D	N	A	AS
13. If you had individual meetings with your teacher: Meeting with the teachers on a one-to-one basis was valuable.	DS	D	N	A	AS
Written papers were graded on:					
14. Technical control (grammar)	DS	D	N	A	AS
15. Organization	DS	D	N	A	AS
16. Development	DS	D	N	A	AS
17. Neatness	DS	D	N	A	AS
18. Insight	DS	D	N	A	AS
19. Style	DS	D	N	A	AS
20. Please list criteria in addition to these used to grade your written papers.					

21. Which one of the above grading criteria was most important? 14 15 16 17 18 19 20

22. Check the title(s) of the texts which were furnished to you for your individual use.

_____ Bain, Beaty, and Hunter, **Norton Introduction to Literature**

_____ Barnet, Berman, & Burto, **Introduction to Literature**

_____ Conley & Cherry, **Poems for Comparison and Contrast**

_____ Roberts, **Writing Themes About Literature**

_____ Other (Did you use only handouts?) _____

	Disagree Strongly	Disagree	Neutral	Agree	Agree Strongly
(For those titles you checked only) I used the above book(s) very little	DS	D	N	A	AS
23. The SUPA Student Manual for the English course was useless in the poetry mini-course.	DS	D	N	A	AS
24. I had a syllabus for the poetry course which listed the assignments.	DS	D	N	A	AS
The following elements were emphasized in the poetry course.					
25. Structure	DS	D	N	A	AS
26. Imagery	DS	D	N	A	AS
27. Tone	DS	D	N	A	AS
28. Historical setting	DS	D	N	A	AS
29. Persona	DS	D	N	A	AS
30. Authors' biographies	DS	D	N	A	AS
31. Other elements.					
32. My teacher often read poetry aloud to the class.	DS	D	N	A	AS
33. Learning to write poetry was part of the course.	DS	D	N	A	AS
34. My first major graded paper was written out-of-class.	DS	D	N	A	AS

35. The approximate length, in number of words, of the last out-of-class essay was: 100 200 300 400 500 600 700 800 900 1000

36. What was the length, in weeks, for your poetry course? 1 2 3 4 5 6 7 8 9 10

37. The course was long enough for the work required. DS D N A AS

38. How many poems, counting both those read in and out of class, were you required to read? (approximate number) _____

Please return to Syracuse University Project Advance, 759 Ostrom Ave., Syracuse, N.Y. 13210.
The enclosed envelope requires no postage.

Thank you

Evaluating the Implementation of an Instructional Design Across Settings

Thomas C. Reeves
Robert E. Holloway

Project Advance
Center for Instructional Development
Syracuse University

A Paper Presented at the Annual Meeting of the
American Educational Research Association
San Francisco, April 9, 1979, Session 8.19

Mailing Address:
Thomas C. Reeves or
Robert E. Holloway
Syracuse University,
Project Advance
759 Ostrom Avenue
Syracuse, NY 13210

This was prepared as part of a larger study which will be available in part in dissertations, abstracts (Reeves) and an NIE funded study (Holloway) which will be available through ERIC in 1980.

ABSTRACT

The purpose of this study was twofold: to develop measures of the implementation of an instructional program and to use these measures to reveal relationships between degrees of implementation of various dimensions of an instructional program and degress of outcomes of the program.

Student and teacher questionnaires were developed as measures of the implementation of the Personalized System of Instruction (PSI) design as used in Project Advance psychology and calculus courses. The student questionnaire was completed by a non-random sample of 708 students in 40 classes in 33 high schools at the end of the spring 1978 semester. The teacher questionnaire was completed by a non-random sample of 34 of the 41 teachers whose classes participated in this study. Additional data collected included: student ratings of teacher and course characteristics, self-reported grade-point-averages and Scholastic Aptitude Test scores, pre- and post-course measures of student attitudes toward the courses, and course achievement indicators (total points earned in psychology, number of hours of credit earned in calculus).

Eight research questions were addressed in a controlled correlational study. Before analyses related to the major questions, the student questionnaire was submitted to a principal components factor analysis with varimax rotation, and six factor scales were developed, labeled, and retained.

Scores from the six scales (Lectures, Mastery/Difficulty, Self-Pacing, Written Materials, Feedback, and Peer Proctors) were also summed to constitute a Student Implementation Rating (SIR). The logical scales from the teacher questionnaire were reconstituted into six subscales supported through item analysis data. Scores from these six subscales (Mastery, Self-Pacing, Written Materials, Lectures as Motivators, Peer Proctors, and Feedback) were also summed to form a Teacher Implementation Rating (TIR).

The differences in SIR and TIR total scale scores among the classes in both courses were statistically significant ($p < .01$). Commonality analysis and partial correlation procedures were used to determine the amount of variance in student achievement accounted for by variation in the implementation of PSI as measured by the SIR when the effects of aptitude and teacher characteristics were controlled. Using the student as the unit of analysis, SIR predicted 1% of the variance in total points in the psychology course and 7% of the variance in the number of credit hours earned in the calculus course ($p < .01$). Using the class as the unit of analysis, SIR predicted 13% of the variance in total points and 48% of the variance in number of credit hours earned ($p < .05$).

Partial correlation analysis was used to determine the amount of variance in student achievement accounted for by variation in the implementation of PSI as measured by the TIR when the effects of aptitude and teacher characteristics

were controlled. No statistically significant relationships ($p < .05$) were detected between TIR and total points in the psychology classes. In calculus, TIR predicted 52% of the variance in number of credit hours earned when the effects of aptitude and teacher characteristics were partialed out ($p < .05$).

Contamination of the pre-course attitude data precluded adequate analyses related to the relationship between implementation and student attitudes.

The SIR and TIR were found to have sufficient reliability (internal consistency and stability) to be used as measures of the implementation of PSI. Arguments and evidence for the content and predictive validity of the measures were presented.

The SIR was judged more effective than the TIR in predicting statistically significant amounts of achievement variance in the psychology and calculus courses. Neither questionnaire was effective in determining the relative contributions of the major dimensions of PSI to the prediction of student achievement. Given the greater effectiveness of the student questionnaire, the relatively greater costs of administering it appear justifiable.

ABSTRACT

The Development and Use of a Measure of the Implementation of the PSI Instructional Design

Thomas C. Reeves
Syracuse University
Center for Instructional Development
115 College Place
Syracuse, N.Y. 13210

with

Robert E. Holloway
Director
Project Advance
759 Ostrom Avenue
Syracuse, N.Y. 13210

The purpose of this study was twofold: to develop a measure of the implementation of the PSI instructional design, and actually to use the measure to reveal the relationships between degrees of implementation of various dimensions of PSI and degrees of achievement on various criterion dimensions of specific PSI courses.

More than 800 students and 30 teachers from two college courses using the PSI design (psychology and calculus) were surveyed and/or interviewed in this study in 1978. All subjects were participants in Project Advance, a cooperative program between Syracuse University and New York State school districts.

Student ratings of the level of use of PSI components were more effective than teacher ratings in predicting statistically significant amounts of achievement variance. The measures did not significantly relate to achievement as measured by grade. The measures did significantly relate the number of credit hours earned in the calculus course ($p < .05$) using classrooms as the unit of analysis.

A Paper Presented at The Fifth Annual Conference of the
Personalized System of Instruction
Washington, D.C.
May, 1979

DIVISION 7: INSTCOM
CHOLULA #7
(ENGLISH)

A METHODOLOGY FOR ASSESSING THE IMPLEMENTATION OF EDUCATIONAL INNOVATIONS: ANALYSIS AND REVISION

Robert E. Holloway
Teachers College,
Columbia University
New York, NY 10027

Thomas C. Reeves
Title XX Training Unit
New York State Office
of Mental Health

This study is an analysis of two surveys developed to measure the implementation of innovations. This line of research is designed to (1) identify generalized qualitative characteristics of prescribed instruction (e.g., autotutorial, personalized systems of instruction), as perceived by learners, (2) describe procedures for developing survey measures to quantify these characteristics and to ultimately, (3) measure the degree to which an instructional procedure has been implemented from the learners' viewpoint.

Methodological problems of measuring the degree of implementation of educational innovations have been extensively discussed in program evaluation literature (cf., Wispe, 1951; Stake, 1967; Cooley & Lohnes, 1976; Gephart, 1976; Fullan & Pomfret, 1977; Hall, 1977). The continuing investigations stem from difficulty in empirically describing treatments in field research. The unmeasured variance in treatments confounds the analysis of treatment-outcomes data.

The first survey in this study was developed using the basic elements of Personalized System of Instruction-designed courses in psychology and calculus. The items of the second survey, used in the poetry unit in an introductory English course, were constructed using the rubrics "materials, structure, behavior, knowledge, and values" (Fullan & Pomfret, 1977).

The findings suggest that critical factors of instruction are either poorly described in course documentation or poorly operationalized. The correlations between the items describing the PSI courses and student grades were not significant. The lack of variance in grades in mastery courses makes this a probable outcome.* Uniform implementation, while not supported by field observations, could be another consideration.

The factor analyses of the English survey suggests that learners may not perceive characteristics of courses as they are conceptualized by curriculum theorists. Recommendations for alternative techniques to further this line of inquiry conclude the paper.

*Correlations between student perceptions of the calculus course and amount of work done as measured by credit earned were significant ($p < .05$).

5/80

APPENDIX C

Project Abstract Format

NIE-G-78-0063

NATIONAL INSTITUTE OF EDUCATION
Basic Skills Group

INVESTIGATOR'S PROJECT ABSTRACT

TITLE: A METHODOLOGY FOR ASSESSING THE IMPLEMENTATION OF EDUCATIONAL INNOVATIONS: ANALYSIS AND REVISION

E Small Grant
Selected Area

INVESTIGATOR(s): Robert E. Holloway, Thomas C. Reeves, Don P. Ely

BUSINESS ADDRESS AND TELEPHONE: Ctr. for Instructional Development, Syracuse University
759 Ostrom Avenue, Syracuse, N.Y. 13210 - (315) 423-2404

PROJECT DURATION: 7/1/78 - 6/30/79

TOTAL FUNDS REQUESTED: $9,744.00

Abstract

This study proposes an analysis of multiple measures used in an assessment of innovation fidelity and the subsequent revision of instruments and methodology. The purposes are to provide valid and reliable methods of measuring critical elements of an innovation and identifying levels of implementation.

Methodological problems of measuring the degree of implementation of educational innovations have been extensively discussed in program evaluation literature (cf., Wispe, 1951; Stake, 1967; Cooley and Lohnes, 1976; Gephart, 1976; Fullan and Pomfret, 1977; Hall, 1977). The continuing investigations stem from difficulty in empirically describing treatments in field research. The unmeasured variance in treatments confounds the analysis of treatment-outcomes data. Thus, administrators and instructional developers must make formative and summative judgement of the effectiveness of programs with little or inaccurate objective information on attributable program effects.

The data will be collected in the spring of 1978. Multiple process measures (observations, inventories, questionnaires) will be used in 65 high schools (appendix) for 2500 senior students, and 120 teachers. The students are predominately white middle class with the sexes equally represented.

Basic psychometrics, such as factoring, item analysis and reliability checks, will be conducted. Canonical analysis will be used to determine the proportion of criterion variance explained by each domain of predictor variables. The degree of relationship between process measures and achievement and attitudes will be used to select critical measures. The basic psychometric analyses will be used to refine those measures retained. The result will contribute to the development of a generalized methodology for measuring process as it relates to outcomes.

SYRACUSE UNIVERSITY
PROJECT ADVANCE
759 OSTROM AVENUE | SYRACUSE, NEW YORK 13210
315/423-2404

We have spent much of your time and ours during visits and seminars specifying the methods and materials used in SUPA courses. To represent the program credibly to other colleges, we have collected considerable evaluation data describing course and program outcomes.

We have also obtained financing from the National Institute of Education for a small study to measure the effects and value of our efforts to achieve comparability between the on- and off-campus Poetry units. The Poetry course has had several other assessments collected this year for a separate study, so we are trying to take advantage of that descriptive information to test the reliability and validity of the enclosed questionnaire.

Since the questionnaire is still in the development stage, the results will not be reported to you nor will records be kept by school, class, or student, except for a follow-up mailing if forms are not returned. A summary of the analysis of the questionnaire for all classrooms combined, however, will be available.

An extra copy of the questionnaire is enclosed for you to complete. Your editorial comments, salacious or otherwise, are welcome.

I appreciate your cooperation and help in passing out the questionnaire and return envelope to students who were in your Poetry class.

Sincerely,

Robert E. Holloway
Director, Project Advance
Center for Instructional
Development

REH:jb
Enc.